图书情报工作
第 65 卷 第 2 期 2021 年 1 月

中文学术图书的 Altmetrics 数据生成动机及数据可用性探析*

——以豆瓣读书平台“短评”为例

■ 魏佳文 韩毅
西南大学计算机与信息科学学院 重庆 400715

摘 要：[目的/意义]探讨用户在“豆瓣读书”发布短评的动机，尝试判断短评作为图书影响力的评估指标是否有效，以期为中文学术图书评价的 Altmetrics 指标建构提供参考。[方法/过程]以 CBKCI 统计报告公布的被引前十的古籍类中文学术图书为样本，采用扎根理论对样本图书在“豆瓣读书”网的 1 142 条“短评”进行分析，探究短评的产生动机及相应短评数据的可用性。[结果/结论]用户产生短评的动机包括讨论、分享、消遣娱乐、外界压力、情感宣泄和资料获取 6 个主范畴。其中，“讨论”动机是最主要的动机，其两个子动机“表达与图书有关的个人经验或想法”和“表达对图书的见解或个人解释”占所有短评数据的四成以上，表明用户不是仅复制图书部分句子或章节，而是试图找到并传播图书最有价值的部分，此类短评数据可用以评价中文学术图书价值；而“分享”“消遣娱乐”“外界压力”“情感宣泄”和“资料获取”等动机相对来说难以较好地反映用户对图书的态度和认知，不宜纳入中文学术图书 Altmetrics 评价指标中。

关键词：中文学术图书 Altmetrics 数据 数据生成动机 数据可用性 豆瓣读书平台 短评
分类号：G250.2
DOI：10.13266/j.issn.0252－3116.2021.02.010

1 引言

图书作为系统化集成化的高密度、长篇幅知识呈现形式，是学术交流体系的重要组成部分，对其进行评价是促进图书出版和学术生态系统健康持续发展的重要环节[1]。图书评价不仅需要从正式交流角度评价其学术影响力，还需要从非正式交流视角评价其社会影响力[2]。Altmetrics 是 Web 2.0 环境下涌现出的一种评价方法，其指标数据主要来源于社交媒体，为衡量学术成果的社会影响力提供了新的视角[3]。

尽管 Altmetrics 蕴含巨大价值，但其可靠性、操作性和适用性仍存在争议，如在“In Metrics We Trust?”研讨会中学者们就曾讨论过 Altmetrics 的价值、潜在作用和研究评估中无意识的不当操作结果等问题[4]。与传统计量学的数据相比，非正式交流中的 Altmetrics 数据是否能在评价中经得住质疑还需要时间的检验。

与传统引用指标可用性受到引用动机影响一样，Altmetrics 指标数据的产生动机也会对指标的价值效度造成直接影响[5]，因而只有深入了解参与科研过程的群体在社交媒体开放空间中存取数据的目的，才能更加准确地评断 Altmetrics 指标在科研评价中的作用。因此，本文将以中文学术图书为研究对象，尝试对 Altmetrics 短评数据产生动机进行深入分析，以期为中文学术图书评价的 Altmetrics 指标构建提供参考。

2 相关研究

2.1 Altmetrics 数据生成动机

Altmetrics 数据生成者是 Altmetrics 数据的源头，理解他们创建 Altmetrics 数据的动机是了解 Altmetrics 指标的基石[6]。已有学者对大众社交媒体 Facebook 上的用户转发和点赞行为、学术社交媒体如 Mendeley 上的用户保存和收藏行为以及用户发布推文行为的动机进行研究，探讨了从众动机[7]、获利动机[8]、社群参与动机[9]、兴趣动机[10]等。深入 Altmetrics 数据的产生

* 本文系四川省社会科学规划重大课题“中文图书 Altmetrics 评价及其在四川哲学社会科学学科影响力分析中的应用”（项目编号：SC17EZD035）研究成果之一。
作者简介：魏佳文（ORCID：0000－0002－1960－319X），硕士研究生；韩毅（ORCID：0000－0001－7021－3229），教授，博士，通讯作者，E-mail：hanyi72@swu.edu.cn。
收稿日期：2020－06－07 修回日期：2020－09－21 本文起止页码：98－106 本文责任编辑：杜杏叶

情境,了解数据背后的用户动机是当前的重要研究内容,但不同学者在不同研究中对动机的分类标准存在差异。

针对 Twitter 用户发布推文动机,T. H. H. Htoo 等[11]将其分为讨论、分享、宣传和获取四类,在此基础上细分出 15 个子动机。J. C. Na[12]借鉴 H. Shema 等[13]的博客引用动机分类将心理学领域用户的英语科学推文动机分为讨论、批评、推荐、触发、扩展、自我推广、辩论、提供数据、伦理以及其他十个类别。K. Schnitzlerk 等[14]在对相关研究[15-18]分析的基础上,将用户为推文添加标签的动机总结为:与有共同兴趣的用户交换意见、获得新粉丝并使推文更容易得到转发等;将发布推文时@其他人动机总结为确保特定用户能接收到推文并鼓励其进行转发;将收藏推文的动机总结为感兴趣等。D. Boyd 等[19]发现用户转发推文是为了通过转发来保存该推文,以备将来个人访问。E. Mohammadi 等[20]对 860 名 Mendeley 用户的调查发现,用户为文档添加标签的动机有预引用、更新知识储备、用于教学活动等。S. Y. Syn 等[21]发现 Facebook 用户和 Twitter 用户转发文献主要动机是为他人提供信息、学习新事物、支持社群利益、享受等。Y. K. Huang 等[22]通过定性与定量相结合的多元分析,发现用户在互联网发布书评主要出于分享、发泄消极情绪、社交互动以及产品参与的动机。张辉等[23]基于使用与满足理论将博客评论动机分为社区参与、情感交流、娱乐及消磨时间。H. Yu 等[24]对新浪微博 Altmetrics 指标前 1%论文做了内容分析,将用户发布科学微博的动机大致分为 4 个方面:提供肤浅介绍传播论文、突出或详细介绍论文学术用途、讨论论文实际用途、突出令人吃惊部分或论文有趣部分等。

对于大众社交媒体 Facebook 上的用户转发和点赞动机,K. Baek 等[25]将 Facebook 用户转发动机分为信息共享、方便和娱乐、打发时间、人际效用、控制和推广工作 6 类,并针对 Altmetrics 动机的理论研究解释了 Altmetrics 数据的产生机制。S. Haustein 等[26]从引用行为理论和社交行为理论角度分析社交媒体行为,对 Altmetrics 本质进行解读;C. Y. Chin 等[27]基于动机理论探究 743 名 Facebook 用户的点赞动机,大致将其分为享乐动机、功利动机、遵从动机、从众动机和归属动机。

关于学术社交媒体 Mendeley 等上的用户保存和收藏动机研究,J. Jiangs 等[28]将 Mendeley 中用户对文献进行标记和将文献加入讨论组的动机分为共享、交流、构建合作网络三类。

2.2 生成动机对 Altmetrics 数据可用性的作用与影响

自 2010 年 Altmetrcis 被正式提出以来,相应数据是否可应用及如何应用于学术评价受到学者们持续关注。学者们[21-22]普遍认为,理解 Altmetrcis 指标内涵并判断数据可用性最直接的途径是探究 Altmetrcis 数据的生成动机,即用户出于何种目的在社交媒体平台上留下科研活动轨迹,可使用的方法包括内容分析法[29]、访谈及问卷调查法[30-31]等。

王鹏飞等[32]对新浪微博进行内容分析,发现大部分为转发、简单转述论文内容,表明微博作者对论文只是一般性关注,没有提供更多的信息;同时也发现有一些微博反映了微博作者对论文的评价、反思和应用,这些微博是对科研目的和作用的深层次传递。S. Lemke 等[33]对研究人员在 18 种社交媒体平台的使用目的进行问卷调查,发现相比于学术博客和学术性网络上对学术成果的提及,脸书和推特上的提及更大程度上反映的是公众传播。

以 Twitter 指标为对象,对 Altmetrics 数据可用性进行探讨的研究相对较多。M. Thelwall 等[34]针对小样本科学推文的内容分析,发现大部分推文主要阐述文献的表征信息,没有对文献内容进行解释,且提及量主要是由于自动转发引起,所以作者认为此类数据几乎无法反映文献的学术价值。L. Bornmann[35]对 Hirsch 发表于 2005 年的 H 指数论文的相关推文进行内容分析,发现其中 37 篇属于简单的转推、转述,6 篇简单提及 H 指数,2 篇谈及个人 H 指数,还有 2 篇是在使用其他计量指标中提及 H 指数,仅有 4 篇对文章的细节进行了讨论并发表了自己的观点。J. C. Na[12]通过分析推文发现,53%推文引用论文总结中的科学发现,12%推文仅包含论文的标题和链接,19%推文是转推,大多数科学推文的作者对所推论文呈中立态度,即大部分推文并没有传达对于科研文献的深入批判性思考和讨论,不能直接反映文献的学术价值。刘晓娟等[36]提炼出 19 种用户动机,引入用户行为理论解释和分析用户动机体现的指标价值,并通过内容分析探讨 Twitter 提及量的应用价值,发现并不是所有动机产生的 Altmetrics 指标数据在科研评价中都具有高应用价值,其中从众动机、消遣动机、无意动机、简化表达、被动动机、社群参与动机等 6 种用户动机反而会产生负向影响,降低指标的应用价值,这在一定程度上回答了 Altmetrics 指标可用性问题。

综上所述,针对 Altmetrics 数据生成动机研究逐渐兴起,分析数据产生动机有助于更加细粒度认识 Altmetrcis 指标内涵,也可以进一步判断数据作为科研评价指标是否有效,但不同学者对动机的分类标准存在差异。多数研究针对用户在社交媒体等各类 Web 应用上对文献的某种操作动机进行探讨,如大众社交媒体 Facebook 上的用户转发和点赞、学术社交媒体如 Mendeley 上的用户保存和收藏以及 Twitter 用户发布推文等,而关于中文学术图书 Altmetrics 数据生产动机的系统性研究明显不足,也鲜有理论支撑。中文学术图书的 Altmetrics 数据在社交媒体上是如何生成的?数据生成动机是什么? 中文学术图书 Altmetrics 数据是否可用于影响力评价? 这些问题还有待进一步解决。

基于此,本文尝试以豆瓣读书平台为例,对中文学术图书 Altmetrcis 数据产生动机做初探性研究,聚焦用户评论对图书本身的学术价值与社会价值产生的作用进行分析。

3 研究设计

3.1 研究方法

本研究主要遵循扎根理论的研究策略[37],对用户在豆瓣读书平台发布短评数据的动机进行分析。扎根理论是通过系统化程序,针对某一现象归纳并引导出结果的一种定性研究方法[38],本研究将采用该理论对豆瓣读书平台短评数据的产生动机进行分析,通过资料采集、分析、编码和理论建模,最终从理论层次上描述现象的本质和意义。

3.2 数据收集

《中文学术图书引文索引》(*Chinese Book Citation Index*,CBKCI)有严格的学术图书遴选流程,其内容标准之一是"原创性、高水平的学术专著"[39]。2017 年底 CBKCI 统计报告发布了"被引排名前十位古籍图书""被引排名前十位近代图书""被引排名前十位现代图书""被引排名前十位外文/译文图书"四大类中文学术图书及其被引频次。经前期检索发现,相较于近代类、现代类以及外文/译文类这三类"被引排名前十位图书",古籍类"被引排名前十位图书"在豆瓣读书平台累积的短评数量较丰富,因此本文选择 CBKCI 统计报告 2017 年底公布的古籍类"被引排名前十位图书"[40]作为研究样本。

豆瓣读书是目前国内较为专业的最具影响力的主流图书阅读分享网站,在知名度、用户活跃度、学术图书覆盖面、内容成熟性、跨平台合作等方面都比较适合作为图书 Altmetrics 指标数据源[41]。豆瓣注册用户可写作、分享短评,短评字数限制在 350 字以内,简短感想、吐槽、推荐皆可。豆瓣活跃用户大多都是阅读爱好者,对图书阅读具有较好的历史积累,对社区具有较高的忠诚度,因此评论大多具有较好的专业性和可参考性[42]。

表 1 CBKCI 报告图书样本集及其被引频次

编号	书名	被引频次/次	编号	书名	被引频次/次
B1	史记	749	B6	左传	508
B2	论语	743	B7	礼记	502
B3	汉书	722	B8	荀子	501
B4	孟子	648	B9	庄子	461
B5	后汉书(全十二册)	545	B10	宋史	444

因此,本研究采用"八爪鱼采集器"从豆瓣读书平台搜集样本图书的短评数据,数据收集结果截至 2019 年 12 月 28 日,如表 2 所示:

表 2 "豆瓣读书"短评数量

编号	书名	短评数量/条	编号	书名	短评数量/条
B1	史记	204	B6	左传	29
B2	论语	255	B7	礼记	22
B3	汉书	123	B8	荀子	81
B4	孟子	188	B9	庄子	65
B5	后汉书(全十二册)	107	B10	宋史	68

3.3 短评产生动机分析

3.3.1 编码

(1)开放编码。开放编码是一种把资料记录及抽象出来的概念"打破""揉碎"并重新综合的过程,其目的是将资料集逐步进行概念化和范畴化[43]。根据开放编码的规则,本文对豆瓣读书平台收集的 1142 条短评文本数据进行编码,通过对标签进行反复分类整理最终归纳出 14 个范畴,编码过程示例见表 3。

(2)主轴编码。主轴编码是对从开放编码中获得的概念和范畴进行关联性分析,找到其中的共同属性并进行分类,发掘并构建范畴之间的相互关系,得出主范畴。主轴编码是开放编码范畴的进一步发展,可以通过整理条件、脉络、策略和结果等帮助了解主范畴的副范畴。对开放编码形成的 14 个概念化范畴进一步组合、琢磨,对概念化内容加以提炼并予以命名,最终得到 6 个主范畴:情感宣泄、消遣娱乐、外界压力、讨论、分享、资料获取。

(3)选择编码。选择编码是不断发掘范畴之间的相互关系,提高概念抽象的层次,定义核心范畴,使其

表 3　开放式编码形成的概念范畴举例

编号	开放编码	概念	部分短评内容
1	图书见解	表达对图书的见解或个人解释	"应该叫孟子的治国思想,仁政本身是完美的,不过没有与之对应的方法论,抑或有向仁发展的君王在实践过程中可能会制造无心伤民之策,没有论语的宽广"
2	图书经验	表达与图书有关的个人经验或想法	"看了《孔子》,马上翻了史记,觉得太不过瘾,就又翻了论语。这才突然想起来,高二的时候选方仁工的选修课就是论语,不过不是为了长知识,而是方便逃课睡觉做物理之类的。而今读来,夫子循循然善诱人,博我以文,约我以礼,欲罢不能。既竭吾才,如有所立卓尔"
3	图书喜好	表达个人喜好、赞成或推荐图书	"史裁如范,千古几人耳？此生最爱的一部书"
4	图书思考	提出一个值得思考的问题	"荀子的思想到底算不算儒家？或者说我们现在所习以为常的儒家学说到底是否能够代表儒家的本来面目？这是一个非常值得思考的问题……"
5	图书讨论	邀请进一步讨论	"是想知道《王莽传》是怎么写的。没读完"
6	图书质疑	对图书的全部或部分进行批评或质疑	"简体横排无注释。《汉书》和《史记》有一些篇章是类似的,但总感觉这些内容相同的篇章中,《史记》读起来更加明白晓畅一些,《汉书》一些地方的遣词用字令人不解。有待研究"
7	图书内容	强调图书的某句话、某些人或某些章节	"井蛙不可以语于海者,拘于虚也;夏虫不可以语于冰者,笃于时也;曲士不可以语于道者,束于教也"
8	图书介绍	简单分享图书或通过肤浅的介绍去传播图书	"古籍(必读)"
9	图书分享	分享给特定的朋友或者将图书指向他们的追随者。	"高考必读经典"
10	负面情绪	当用户意识到阅读的图书不如他们想象或期望的那样好	"读了跟没读一样"
11	被迫阅读	被迫阅读(工作或论文需要)	"为了工作我看了 1.2% 已经不容易了～"
12	打发时间	用户无事可做或仅仅将短评作为他们打发时间的一个方式	"无需赘言"
13	休闲娱乐	出于放松和休闲的目的而评论,通过评论过程来消磨时间	"在等机的时候看的"
14	获取资料	获取与图书相关的资料或者其他版本的图书	"想找一个装帧素雅没有注释的繁体版本,班氏真是太啰嗦了"

包含其他范畴,并思考一个纳入所有范畴、关系的故事线,扼要说明全部现象的核心[44]。本研究主轴编码阶段已抽象出 6 个主范畴,通过维度、属性等方面的进一步分析,最终提炼了用户发表短评的个体相关动机、图书相关动机以及利益驱动动机 3 个核心范畴,如表 4 所示:

表 4　主轴编码和选择编码

选择编码	主轴编码	开放编码	频次	百分比(%)
个人相关	消遣娱乐	打发时间	185	16.20
		休闲娱乐	20	1.75
	外界压力	被迫阅读	34	2.98
图书相关	情感宣泄	负面情绪	14	1.23
	分享	图书内容	152	13.31
		图书介绍	71	6.22
		图书分享	22	1.93
	讨论	图书见解	124	10.86
		图书经验	337	29.51
		图书喜好	94	8.23
		图书思考	23	2.01
		图书讨论	11	0.96
		图书质疑	49	4.29
利益驱动	资料获取	获取资料	6	0.53

3.3.2　理论饱和度检验

理论饱和度(Theoretical Saturation)是指即便进行额外抽样也不会再发现新的概念范畴,也不会再生成新的理论。随机抽取豆瓣读书平台其他中文学术图书短评内容进行编码分析以检验理论饱和度,发现样本资料所提取的概念均已被现有范畴包含,没有获得新范畴,因此认为理论已经达到饱和。

4　结果分析与讨论

4.1　结果分析

本研究对豆瓣读书平台中 10 部高被引中文学术图书的 1 142 条短评进行扎根理论分析,将用户发布短评的动机归纳为个人相关、图书相关以及利益驱动三大类。个人相关动机包括消遣娱乐和外界压力,图书相关动机包括情感宣泄、分享和讨论,利益驱动动机即资料获取。

(1)消遣娱乐。娱乐是指出于放松和休闲的目的而评论图书。用户在豆瓣平台阅读图书可以满足他们的兴趣和好奇心、消磨时间,获得乐趣。消遣娱乐是个体自发阅读图书后发布评论,是个体内在需要激发阅读图书和发布评论行为意愿而产生的动机。该动机在张辉[23]、K. Beak[24]、C. Y. Chin[28]中也得到了广泛的关

注。在本研究中，其涉及的主要维度是打发时间和休闲娱乐。打发时间正向影响用户发布短评，指用户无事可做或仅仅将短评作为他们打发时间的一个方式；休闲娱乐也正向影响用户发布短评，指用户出于放松和休闲的目的而评论，通过评论过程来消磨时间。在消遣娱乐中，“打发时间”的动机占总短评的 16.2%，这类短评极具主观性、随意性，质量参差不齐。

（2）外界压力。与消遣娱乐相反，外界压力是指个体在被迫阅读图书的背景下发布评论，用户通过发布短评来宣泄自己被迫阅读的不满情绪。该范畴主要用被迫阅读来解释，主要因为工作或论文需要不得不阅读图书。与该类动机相类似的是在刘晓娟等[36]研究中提到的“被动”动机，用户受外界因素压制，迫不得已评论。例如“为了工作我看了 1.2% 已经不容易了”，就是用户因工作需要去阅读读书，进而发布评论。然而外界压力（占比 2.98%）动机产生的数据无法体现用户的真正态度，反而会使得指标的应用价值大打折扣，由此产生的数据可视为无效数据，不建议用于科研评价中。

（3）情感宣泄。用户在豆瓣平台阅读过程中会产生积极或者消极情绪，通过发表评论能够分享积极情绪和缓解消极情绪。一般地，由图书相关内容引起消极情绪而发布的短评，称为情感宣泄。这种负面情绪可能由于阅读体验不好，或者图书太难、用户无法理解等引起，既与个人的文化程度和阅读能力等相关，又与图书本身的类型、难易、好坏等相关联。对情感宣泄（占比 1.23%）这一范畴主要从用户的负面情绪来理解，当用户意识到图书在阅读后不如他们想象或期望的那样好，或者他们不同意书中的论点以减轻后悔或失望时，他们将通过在豆瓣上发布书评来表达他们不愉快的阅读体验。

（4）分享动机。图书阅读有深浅之分，浅阅读是指浏览式、泛泛的阅读，不用对所阅读的图书内容做深入的思考。在浅阅读语境下，用户主要分享关于图书本身的内容或者将图书简单推荐给适合阅读此书籍的人群，由于未深入思考图书内容，所以此类短评并没有传达或者仅仅只是简单传达自己的观点或解释。对于分享这一范畴，可从图书内容、图书介绍、图书分享三个方面进行理解。用户为了强调图书的某句话、某些人或某些章节而发布短评，即为图书内容，此类短评通常仅仅是复制图书本身的内容，不能体现用户支持或反对的态度；如若用户支持图书，则希望通过发布短评简单分享图书或通过肤浅的介绍去传播图书，即图书介绍；用户也可以将图书分享给特定的朋友或者将图书指向他们的追随者而发布短评，即图书分享动机。“分享”中最常见的动机是“图书内容”，涵盖了 13.31% 的短评；其次是“图书介绍”动机，如“不错”、“还可以”等短评，占 6.22%。与 J. C. Na[12]研究类似，此类动机的大多数短评并没有对图书深度讨论。

（5）讨论动机。与浅阅读相比，深阅读是指进入内容情境，加入用户许多深刻思考的阅读。在深阅读语境下，用户有时是伴随着对历史、社会经验、人生经历一起思考讨论，有时兴奋，有时忧郁，有时赞同，有时质疑。因此，用户发布的“讨论”短评是为表达个人对图书的解释、观点、想法、批评或怀疑，该动机同样涉及与图书相关的情感因素。用户可以为了表达个人喜好、赞成或推荐图书而发布短评，即图书喜好；也可以因为对图书的全部或部分内容进行批评或质疑而发布短评，即图书质疑。此外，用户可能单纯地通过短评表达对图书的见解或个人解释，即图书见解；或者表达与图书有关的个人经验或想法，即图书经验。在图书阅读过程中，用户可能会通过发布短评提出一个值得思考的问题，即图书思考；也可能为了邀请其他读者进一步讨论发布短评，即图书讨论。“讨论”中最常见的动机是“图书经验”，涵盖了 29.51% 的短评，其次是“图书见解”与“图书喜好”，分别占 10.86% 和 8.23% 的比例。这类短评揭示了用户不是仅仅复制图书某句话或某个章节，而是确实试图找到图书最有价值的部分并且对其进行传播。

（6）资料获取。资料获取是指获取与图书相关的资料或者其他版本的图书。资料获取可以发生在图书阅读的任何一阶段，用户受个人相关因素的影响，在自发阅读和被迫阅读驱动下可因利益去获取资料；同时，在浅阅读和深阅读语境中，无论是为了浏览式、泛泛的阅读，还是加入许多深刻思考的阅读，用户都希望使得豆瓣阅读平台的图书质量和服务得到提高，从而发布短评。该短评仅占 0.53%，说明在豆瓣平台上，只有极少数用户是出于获取其他版本图书的目的而发布短评。这一原因在 T. H. H. Htoo[11]的研究中也得到证实，心理学领域用户发布科学推文很少出于“获取（索要全文）”动机。同样，豆瓣用户发布短评往往也不是因为想要获取与图书相关的资料。

4.2 模型构建

仔细分析消遣娱乐、外界压力、情感宣泄、分享、讨论和资料获取这 6 个主范畴可建构如下关系（见图 1）：在一定的图书阅读语境中，在个人相关因素的驱动

下促使用户产生阅读图书的意向，从而引发与图书相关的一系列情感宣泄、分享及讨论行为；用户在阅读过程中，同时出于利益因素考虑去获取图书相关资料，发布短评。其中，个人相关因素既可以成为用户发布短评的直接因素，同时也可以成为图书相关和利益驱动的间接因素，即图书相关和利益驱动会受内在需求因素的影响进而影响用户最终发布短评的动机。

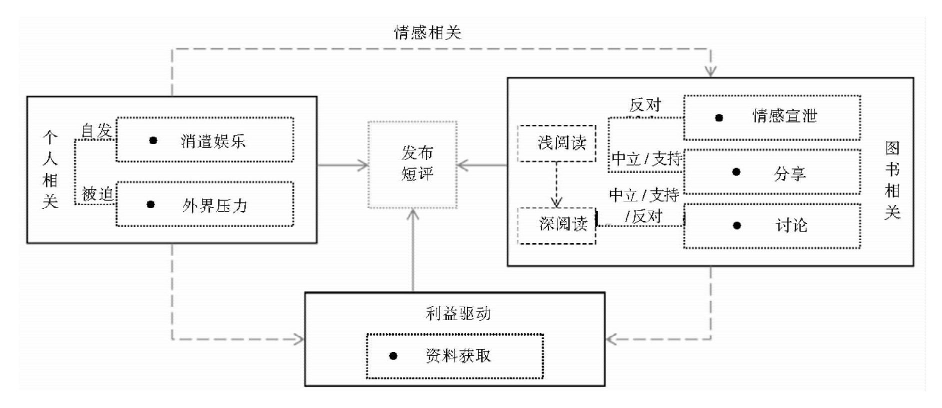


图 1　“豆瓣读书”平台用户发布短评动机模型

消遣娱乐和外界压力是促成用户发布短评的个人情感相关因素。用户在豆瓣平台阅读过程中受外界压力被迫阅读，通过发表评论能够缓解这种不满情绪；也有用户将在豆瓣读书平台整个阅读过程中发表对图书的体验、评价等作为一种娱乐享受，即消遣娱乐动机。这两类动机都涉及用户在阅读过程中的情感体验，主要与用户阅读图书的目的有关，而不是由图书相关内容引发；而出于讨论、分享和情感宣泄目的发布的短评，则可以很好地反映用户对图书的态度和认知。其中讨论范畴，既包含用户对图书表示支持的短评，如表达个人喜好、赞成或推荐图书类动机，也涉及用户通过发表评论来表达给自己带来不愉快阅读体验的短评，如表达对图书的全部或部分进行批评或质疑类动机。这些与图书相关的短评既能体现用户对图书表示支持或反对的情感态度，也包含着用户对图书内容的理解与体验。另外，在浅阅读语境中，虽然用户未对图书内容加以深刻思考，但通过简单阅读后同样可表明自己的支持或反对态度，即是由图书相关内容引发用户的积极（分享）或消极情绪（情感宣泄）；此外，无论用户出于哪种目的开展图书阅读，或是在阅读中处于哪种阶段，用户都希望通过发表评论，使得豆瓣阅读平台的图书质量和服务质量得到提高、平台的服务流程更加完善，即利益驱动也是促进用户发布短评的动机。

4.3　结果讨论

从数据编码总体结果来看，“讨论”作为豆瓣读书用户发布短评最主要的动机，占比达 55.87%，“分享”和“消遣娱乐”次之，“资料获取”最少。由此可见，用户在豆瓣读书平台发布短评时，最主要的目的是表达见解和交流讨论，这也体现了豆瓣读书平台的性质。豆瓣读书是专业的社会化阅读服务平台，鼓励用户创作内容、发表评论，已为平台凝聚了相当的人气，对于用户社会化阅读交流和讨论起到了促进作用[45]。因此，用户为了交流讨论而发布短评也不足为奇。其次，21.45% 的短评是“分享”，即转发、简单转述图书内容，没有明确表示赞成或质疑，具体的用户动机难以判断，因此这类动机不适合作为评价中文学术图书影响力的指标，但可以在一定程度上体现用户对图书的一般性关注。“消遣娱乐”动机的短评占比 17.95%，这类短评并没有任何有意义的内容，同样在刘晓娟等[36]的研究中也表明，消遣动机（娱乐或消磨时间而引发社交行为）或者无意动机（无目的或无意识的社交行为）产生的数据应用价值很低。此外，还有 1.23%、2.98% 和 0.53% 的短评是出于“情感宣泄”“外界压力”和“资料获取”的动机。

理解用户产生短评数据的动机，是为了进一步评判 Altmetrics 数据是否可用。本研究用“数据可用性”表征用户行为数据与学术成果影响力的契合程度，即数据本身在特定语境下是否适用学术成果影响力评价。参考聂磊等[46]对网络行为数据适用性的定义，围绕以下两个方面讨论 Altmetrics 数据可用性问题：①样本代表性，即数据中的样本能够有效推断估计总体的程度；②测量效度，即基于特定数据进行的测量能在多大程度上反映研究概念的真实含义[47]。

一方面，在样本代表性上，本研究以 CBKCI 统计报告 2017 年公布的古籍类“被引排名前十位图书”作为研究样本，CBKCI 有严格的学术图书遴选流程，是

“对学术图书的学术质量和影响力进行综合分析”的新兴中文学术图书评价工具;其次,本研究数据来源于豆瓣读书平台,该平台是当前运营比较成功、比较成熟的社会化阅读平台,它以图书为中心建立关于图书的阅读社区,体现了数字化网络下环境下用户的社会化阅读取向,它的书评功能为研究社会化阅读平台用户的交互行为提供了丰富的数据资源。因此,样本数据具有代表性。

另一方面,在测量效度上,本研究通过分析短评数据发现,55.87% 的短评包含用户对图书的个人解释、观点、想法、批评和怀疑,21.45% 的短评涉及转发、简单转述图书内容,1.23% 的短评是由图书相关内容引发的情感宣泄,而其余 21.46% 的短评仅是消遣娱乐、外界压力以及资料获取。一半左右的短评用户传递了对图书的深度讨论与见解,这部分短评数据可以反映图书的学术影响力和社会影响力,适用于中文学术图书影响力评价。另外近乎一半的短评数据或简单分享图书,或通过肤浅介绍去传播图书,再或者仅仅出于消遣娱乐、情感宣泄、外界压力以及资料获取的目的,并没有传达对于图书的深入批判性思考和讨论,不能直接反映图书的学术价值,因此这部分短评数据不适用于中文学术图书影响力评价,其相应指标也不应纳入到 Altmetrics 计量范畴。

5 结论与研究展望

在 Altmetrics 提出之初,学者们将其作为一种潜在的影响力评价指标,以期弥补传统引文评价在社会影响力方面的不足。随着研究的不断深入,人们已经意识到推文量、下载量、评论数等指标已无法全部表达 Altmetrics 的含义,因而基于社交媒体平台上 Altmetrics 数据生产动机来鉴别对应数据的计量可用性为 Altmetrics 的深入研究提供了新的路径。

本研究对 1142 条高被引中文学术图书的短评进行扎根理论分析,发现用户在豆瓣读书平台发布短评的动机可以分为讨论、分享、情感宣泄、外界压力、消遣娱乐和资料获取 6 个主范畴。与 T. H. H. Htoo 等[11]和 J. C. Na[12]不同的是,本研究发现“讨论”动机(55.87%)是最主要的动机,其中“表达与图书有关的个人经验或想法”(29.51%)和“表达对图书的见解和个人解释”(10.86%)两个子动机占比达 40.37%,表明将近一半的用户不是以肤浅介绍去评论图书,而是详细介绍图书的价值或与图书相关的个人经验,传递了对图书的深度讨论与见解。“分享”动机(21.45%)次之,其中 13.31% 的用户仅仅转述图书的某句话、某些人或某些章节。此外,“消遣娱乐”动机的短评占比 17.95%,这类短评没有任何计量学意义的内容。同样,还有 1.23%、2.98% 和 0.53% 的短评分别出于“情感宣泄”“外界压力”和“资料获取”的动机,与图书质量及影响没有太多的关系,也不宜纳入 Altmetrics 计量范畴。

因此,如果仅仅以短评次数作为评价指标,会引起很大的评价偏差。因为豆瓣读书网中 55.87% 的短评可用来评价图书,其他短评的可用性还有待商榷。具体而言,基于“讨论”动机的短评,包含了用户对图书的深度思考与认知,这类短评可用来评价中文学术图书质量及相应影响。其他四个动机在所有短评中也占了将近一半的比例,出于“分享”“外界压力”“消遣娱乐”“情感宣泄”“资料获取”目的的短评,并不能很好地反映用户对图书的态度和认知。故这些比较随意的、与图书内容无关的短评,不宜纳入中文学术图书评价指标体系中。

本文仅以豆瓣读书平台的短评数据为例,对中文学术图书 Altmetrics 数据生产动机进行了初步探讨,仍有许多不足之处。首先在数据的获取上,豆瓣设置反爬机制,采用八爪鱼仅获取《史记》和《论语》的部分短评,因此本研究仅代表的是所采集到数据的基本情况,并且获得的样本量较小,结果难免带有片面性。其次,本研究仅以古籍类“被引排名前十位图书”作为样本,没有考虑学科类别对研究结果是否会有影响。此外,本研究运用了扎根理论研究策略,在研究中虽尽可能忠实于原始资料,但限于研究者的编码分析能力,在策略的运用上可能存在不足,需要在后续研究中对本研究结论继续加以验证。

未来的研究中,可以考虑对书评、读书笔记、豆瓣评分等数据进行扩充,同时还可以考虑从时间纵向维度和用户角度对数据生产动机作对比分析,更加深入地研究中文学术图书 Altmetrics 数据生产动机,为中文学术图书 Altmetrics 评价提供坚实的数据可用性基础。

参考文献:

[1] 任红娟. 我国图书评价方法研究述评[J]. 图书情报知识,2016(5):22-29.

[2] 韩毅. 非正式交流回归语境下科技评价的融合路径取向[J]. 中国图书馆学报,2016,42(4):64-74.

[3] PRIEM J, HENNUBGER B H. Scientometrics 2.0:new metrics of scholarly impact on thesocial Web[EB/OL]. [2020-08-27]. http://pear. accc. uic. edu/ojs/index. php/fm/article/viewArticle/2874.

[4] In metrics we trust? Prospects & pitfalls of new research metrics [EB/OL]. [2020－08－27]. http://www.sussex.ac.uk/spru/newsandevents/2014h/conferences/metricsworkshop.
[5] 刘晓娟,赵卓婧,宋婉姿. Altmetrics 在学术评价中的可用性研究综述[J]. 数字图书馆论坛,2017(8):38－45.
[6] 余厚强,别克扎提·木拉提. 从 ISSI 2019 会议解读替代计量学研究新进展[J]. 情报理论与实践,2020,43(7):157－164.
[7] 金燕,翟丽辉. 社交媒体环境下用户转发行为对信息质量影响的调查与分析[J]. 图书馆理论与实践,2016(9):91－95.
[8] BARILAN J. JASIST 2001－2010[J]. Bulletin of the Association for Information Science and Technology, 2012,38(6):24－28.
[9] 孙会,李丽娜. 高频次转发微博的特征及用户转发动机探析－－基于新浪微博"当日转发排行榜"的内容分析[J]. 现代传播(中国传媒大学学报),2012,34(6):137－138.
[10] NEYLON C. Altmetrics: what are they good for? [J/OL]. Plos-opens, 2014. [2020－08－27]. https://blogs.plos.org/opens/.
[11] HTOO T H H, NA J C. Who are tweeting research articles and why? [J]. Journal of information science theory and practice, 2017,5(3):48－60.
[12] NA J C. User motivations for tweeting research articles: a content analysis approach[C]//International conference on asian digital libraries. Seoul: Springer, 2015: 197－208.
[13] SHEMA H, BARILAN J, THELWALL M, et al. How is research blogged? A content analysis approach[J]. Journal of the Association for Information Science and Technology, 2015, 66(6): 1136－1149.
[14] SCHNITZLER K, DAVIES N, ROSS F, et al. Using twitter? to drive research impact: a discussion of strategies, opportunities and challenges[J]. International journal of nursing studies, 2016, 59(7): 15－26.
[15] POWER A. Twitter's potential to enhance professional networking [J]. The British journal of midwifery, 2015, 23(1): 65－67.
[16] FAHLBERG B. Leveraging social media to facilitate recruitment and dissemination in doctoral research[J]. Advances in nursing doctoral education and research, 2015,3(1): 10－14.
[17] SHIELDS R. Following the leader? Network models of "world－class" universities on twitter[J]. Higher education,2016,71(2): 253－268.
[18] O'CONNOR A, JACKSON L,GOLDSMITH L, et al. Can I get a retweet please? health research recruitment and the twitter sphere [J]. Journal of advanced nursing, 2014,70(3): 599－609.
[19] BORD D M, GOLDER S, LOTAN G. Tweet, Tweet, retweet: conversational aspects of retweeting on Twitter[C]//2010 43rd Hawaii international conference on system sciences. Hawaii: IEEE, 2010:1－10.
[20] MOHAMMADI E, THELWALL M, KOUSHA K. Can Mendeley bookmarks reflect readership? a survey of user motivations[J]. Journal of the Association for Information Science & Technology, 2016,67(5):1198－1209.
[21] SYN S Y, OH S. Why do social network site users share information on facebook and twitter[J]. Journal of information science, 2015, 41(5): 553－569.
[22] HUANG Y K, YANG W I. Dissemination motives and effects of internet book reviews[J]. The electronic library,2010,28(6):804－817.
[23] 张辉,徐晓林. 博客评论行为动机因素实证研究[J]. 情报杂志,2013,32(11):107－109,201.
[24] YU H, XU S, XIAO T, et al. Global science discussed in local altmetrics: weibo and its comparison with twitter[J]. Journal of Informetrics, 2017, 11(2): 466－482.
[25] BAEK K, HOLTON A, HARP D, et al. The links that bind: uncovering novel motivations for linking on facebook[J]. Computers in human behavior,2011,27(6):2243－2248.
[26] HAUSTEIN S, BOWMAN T D, COSTAS R. Interpreting "Altmetrics": viewing acts on social media through the lens of citation and social theories[EB/OL]. [2019－12－10]. http://arxiv.org/ftp/arxiv/papers/1502/1502.05701.pdf.
[27] CHIN C Y, LU H P, WU C M. Facebook users' motivation for clicking the "like" button[J]. Social behavior and personality, 2015,43(4):579－592.
[28] JIANG J, NI C, HE D, et al. Mendeley group as a new source of interdisciplinarity study: how do disciplines interact on mendeley? [C]// ACM/IEEE－cs joint conference on digital libraries. Newark: ACM, 2013: 135－138.
[29] SHEMA H, BAR－LLAN J, THELWALL M. How is research blogged? A content analysis approach[J]. Journal of the Association for Information Science & Technology, 2014,66(6):1136－1149.
[30] GHAISANI A P, HANDAYANI P W, MUNAJAT Q. Users' motivation in sharing information on social media[J]. Procedia computer science, 2017, 124:530－535.
[31] SUGIMOTO C R. Theories of informetrics: a festschrift in honor of blaise cronin[M] //HAUSTEIN S, BOWMAN T D, COSTAS R. Interpreting "Altmetrics": viewing acts on social media through the lens of citation and social theories. Berlin: De Gruyter, 2016:372－405.
[32] 王鹏飞,刘烜贞. 基于内容分析的 Altmetrics 本质研究[J]. 图书情报工作,2017,61(2):114－120.
[33] LEMKE S, PETERS I. Coping with Altmetrics' heterogeneity － a survey on social media platforms' usage purposes and target groups for researchers[C]//Proceedings of the 17th International conference on Scientometrics and Informetrics. Rome: Italy, 2019: 2320－2325.
[34] THELWALL M, TSOU A, WEINGART S, et al. Tweeting links to academic articles[J]. Cybermetrics: international journal of Scientometrics, Informetrics and Bibliometrics, 2013,17(1): 1－8.
[35] BORNMANN L. What do altmetrics counts mean? a plea for content analyses[J]. Journal of the Association for Information Science

and Technology, 2016, 67(4): 1016 – 1017.
[36] 刘晓娟,王晨琳. 用户动机视角下的 Altmetrics 指标研究[J]. 图书情报知识,2020(3):63 – 70,96.
[37] DENSCOMBE M. 怎样做好一项研究——小规模社会研究指南[M]. 陶保平,译. 上海:上海教育出版社,2011: 73.
[38] MACMILLAN K, KOENIG T. The wow factor: preconceptions and expectations for data analysis software in qualitative research[J]. Social science computer review, 2004,22(2): 179 – 186.
[39] 项目介绍. 中文学术图书引文索引[EB/OL]. [2020 – 08 – 27]. http://cssrac. nju. edu. cn/cbkci/index. php? m = Home&c = Article&a = lists&category = project – intr.
[40] CBKCI: 被引排名前十位图书[EB/OL]. [2020 – 08 – 27]. http://css - rac. nju. edu. cn/cbkci/index. php? m = Home&c = Article&a = lists&category = stat_2.
[41] 李明,李江,陈铭,等. 中文学术图书引文量与 Altmetrics 指标探索性分析及其启示[J]. 情报学报,2019,38(6):557 – 567.
[42] 赵红. 豆瓣阅读平台研究[D]. 南京:南京大学,2016.
[43] 韩正彪,周鹏. 扎根理论质性研究方法在情报学研究中的应用[J]. 情报理论与实践,2011,34(5):19 – 23.
[44] PANDIT N R. The creation of theory: a recent application of the grounded theory method[J]. The qualitative report, 1996, 2(4): 1 – 14.
[45] 张云. 基于"豆瓣读书"模式的图书馆社会化阅读推广[J]. 晋图学刊,2014(5):1 – 6.
[46] 聂磊,王延飞. 网络行为数据的适用性评估问题初探[J]. 图书情报工作,2019,63(6):29 – 34.
[47] 巴比. 社会研究方法[M]. 邱泽奇,译. 11 版. 北京:华夏出版社,2009.

作者贡献说明:
魏佳文:负责数据收集与分析,以及论文的撰写与修改;
韩毅:负责论文选题,指导并修改完善论文。

Generation Motivation of Altmetrics Data and its Availability to Evaluate Chinese Academic Books: A Case Study of "Short Comments" on Douban Reading Platform

Wei Jiawen　Han Yi
College of Computer and Information Science, Southwest University, Chongqing 400715

Abstract: **[Purpose/significance]** This paper discusses the users' motivation to publish short reviews on the Douban Reading platform, and tries to judge whether the short reviews are effective as indicators to evaluate the influence of books, so as to get inspiration for the indicator construction for Altmetrics to assess Chinese academic books. **[Method/process]** Taking the top 10 cited books of ancient Chinese academic ones published in the CBKCI statistical report as samples, this paper used the grounded theory to analyze the 1,142 "short comments" of the sample books on Douban Reading website, and explored the motivation of the short comments and its availability to assess the social influences of academic books. **[Result/conclusion]** 6 main motivation for users to generate short reviews are identified, such as discussion, sharing entertainment, external pressure, emotional venting, and data acquisition. Among them, the "discussion" motivationis is the main one and its two sub – motivations "expressing personal experiences or ideas related to books" and "expressing opinions or personal explanations on books" account for more than 40% of all short review data. It indicates that most users do not only just copy a certain sentence or a certain chapter of a book, but try to find and spread the most valuable information of the book. Such short – review data can be used to evaluate Chinese academic books. Others motivation, such as "share", "recreational entertainment", "outside pressure", "emotional vent", and "data acquisition", are relatively difficult to reflect well users' attitude and cognition about the books, and these data should not be included in the Altmetrics index to evaluate Chinese academic books.

Keywords: Chinese academic books　Altmetrics data　motivation for data generation　data availability　douban reading platform　short review data